

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/27269010>

Strong stems in the German mental lexicon: Evidence from child language acquisition and adult processing

Article · July 2001

Source: OAI

CITATIONS

13

READS

75

4 authors, including:



Sonja Eisenbeiss

University of Cologne

38 PUBLICATIONS **493** CITATIONS

[SEE PROFILE](#)



Joana Cholin

Bielefeld University

21 PUBLICATIONS **479** CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Psycholinguistics and Linguistic Fieldwork [View project](#)



Syllable processing and syllabic representation [View project](#)

Strong Stems in the German Mental Lexicon: Evidence from Child Language Acquisition and Adult Processing¹

Harald Clahsen (University of Essex), Peter Prüfert (University of Düsseldorf), Sonja Eisenbeiss (Max Planck Institute for Psycholinguistics, Nijmegen), and Joana Cholin (Max Planck Institute for Psycholinguistics, Nijmegen)

Abstract

This study addresses the question of how marked stem forms are represented in the mental lexicon focussing on the marked stems of the so-called strong verbs of German. Results from (i) error analyses of children's spontaneous speech, (ii) elicited production experiments with children and adults, and (iii) word-recognition experiments will be presented to show that inflected word forms that consist of regular affixes and strong stems have decomposed representations and that strong stems constitute subnodes of hierarchically structured entries with underspecified grammatical feature content.

1. Introduction

This paper examines how inflected word forms such as those in (1) are represented in the mental lexicon and how they are acquired by children:

- (1) (ich) werf-e (du) wirf-st (sie) warf-en
 '(I) throw-1ST SG PRES' '(you) throw-2ND SG PRES' '(they) throw-3RD PL PRET'

Morphologists have posited (different kinds of) morphological rules and representations to analyze such forms. The forms in (1), for example, are derivable from combinatorial rules that combine stems or roots with regular affixes. Affixes (or exponents) may be represented in inflectional paradigms defined by their morpho-

¹The research reported in this paper is supported by the *Deutsche Forschungsgemeinschaft* (DFG) 'German Research Council' (grant SFB 282/C7). We thank the members of our research group, Ingrid Sonnenstuhl, Meike Hadler, Helga Weyerts, Kerstin Mauth, and Axel Huth for assistance in designing experimental materials and for many useful ideas. We are also grateful to Dirk Janssen and Claudia Felser for comments and helpful suggestions.

syntactic features. The relationships between the various stem alternants in (1) may either be captured by (morphologically conditioned) phonological rules (see e.g. Beedham 1994), or they may be directly encoded in the lexicon, through lexical redundancy rules (Chomsky 1970, Jackendoff 1975), default inheritance representations with underspecified entries (Corbett & Fraser 1993, Wunderlich 1996), or equivalent mechanisms.

Psycholinguists have asked the additional question of which of these morphological operations and representations are employed by the speaker/hearer in language comprehension and production, and by children in language acquisition. After all, one possibility would be that even though affixation, paradigms, structured lexical entries and the like provide useful tools for the description of patterns among inflected word forms, they do not play any role for mental representation and acquisition. This latter view is indeed held by many psycholinguists working within connectionist approaches to language (see Rumelhart & McClelland 1986, Elman et al. 1996, Bybee 1995, among many others). In these accounts, inflected words are all stored in associative networks which through repeated exposure to multiple sets of inflected words will create connections among the various forms. In this way, the network forms patterns that range over sets of connections, and morphological rules and representations come out as secondary, derivable from associations between words. In English, for example, there are thousands of past tense forms ending in *-ed*, and in a connectionist network this leads to the creation of a strong pattern that makes the network behave as if it had a separate *-ed* rule, even though the network itself does not utilize any kind of morphological rules.

On the other hand, many psycholinguists have argued that morphological structure and representation *is* important for understanding the acquisition and processing of inflected and derived word forms. Several psycholinguistic studies have focused, for example, on the contrast between regular and irregular inflection (Pinker 1999, Marslen-Wilson & Tyler 1998, Clahsen 1999), and there is now evidence from different languages for a dual-mechanism account according to which regulars are segmented or parsed into their constituent morphemes while irregular forms are stored as whole forms in memory. However, inflected word forms do not always fall neatly into one of these two clusters. Consider the forms in (1) again and note that the person and number endings are all highly regular; they are easily segmentable and can be attached to almost any verb. On the other hand, the forms in (1) differ in terms of their stems: *werf-* is the unmarked (or so-called ‘weak’) stem/root form, which is identical to that of the infinitive (*werf-en*), whereas *wirf-* and *warf-* are marked (alternated) stems which are different from that of the infinitive. In the present study, we will examine how different stem forms of the same lexeme are represented in the mental lexicon. One question we will address is whether strong stems such as *wirf-* and *warf-* are rule-governed, e.g. derived from *werf-*, or whether they are lexically stored. Another question concerns the role of morphological features for the representation of stem forms. The phonological string *wirf-*, for example, appears as the stem in both imperatives and 2nd/3rd sg. forms. Does that mean that there are two stem forms *wirf-*,

which are distinguished in terms of morphological features, or is one stem form based on the other - and how are these relationships represented? Psycholinguistic evidence from different sources will be presented that bears on these questions, results from an analysis of strong verb formations in the speech of young children learning German as their native language, and experimental results from elicited production tasks with children and adults. We will also report earlier findings from on-line lexical decision and priming experiments.

Before turning to the empirical results, we will provide a brief description of strong verb inflection in German.

2. Strong verb inflection in German

There are about 160 simplex verbs in German that belong to the strong class. These verbs have marked stems in present tense, preterite, or participle forms. Most of them (= 155) fall into three minor classes, illustrated in (2); see Wunderlich & Fabri (1995) for a detailed classification. While A-B-A and A-B-C verbs have differently marked stem forms in preterite and participle forms, A-B-B verbs exhibit the same vowel change for both preterites and participles. Moreover, a large number of strong verbs have subjunctive forms with unlauded preterite stems, as shown in (2) for *gab-* / *güb-* ‘gave’, *flog-* / *flög-* ‘flew’ and *sang-* / *säng-* ‘sang’. There is also a small number of strong verbs in which 2nd sg. and 3rd sg. present tense forms as well as imperatives have fronted vowels (e.g. *werfen* vs. *er wirft* ‘to throw’ vs. ‘he throws’).

(2)

	Infinitive	Preterite	Subjunctive	Participle	
A-B-A	geben	gab-	gäb-	gegeben	‘to give’
A-B-C	singen	sang-	säng-	gesungen	‘to sing’
A-B-B	fliegen	flog-	flög-	geflogen	‘to fly’

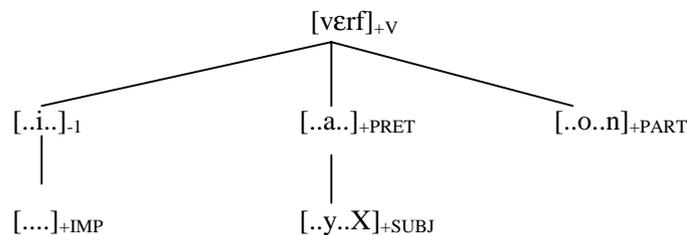
Except for suppletive forms such as those of *sein* ‘to be’ and the forms of modals like *dürfen* ‘to be allowed to’, the marked stems of strong verbs are regularly inflected for person and number in the same way as the unmarked stems of weak verbs.

The linguistic representation of stem alternations is controversial. One suggestion is to derive stem alternants from a single underlying form by (morphologically conditioned) phonological rules. Several attempts have been made to demonstrate this in respect to strong verb formation in German, see Beedham (1994, 1995/1996), Bittner (1996), Barbour (1982), among others. Beedham, for example, observes that particular sequences of onsets and rhymes of root syllables coincide with whether or not a verb has strong stem alternants in German. For instance, verbs whose root syllables have the rhyme *-ing-* (e.g. *singen* ‘to sing’) tend to be strong, whilst verbs with root syllables that have the onset [va] (e.g. *wagen* ‘to dare’) are predominantly

weak. Beedham posits corresponding phonological rules that specify 50 relevant sequences of onsets and/or rhymes to establish the phonological sequences which, in his view, distinguish weak from strong verb roots. One problem of any account that attempts to handle stem alternations by phonological rules is the large number of exceptions and counter-examples; see Durrell (1980, 2000), Wiese (1996).

An alternative solution is to represent stem alternants in the lexicon. One suggestion has been to directly list stem forms such as *wirf-*, *warf-*, *würf-* etc. in the permanent lexicon in much the same way in which suppletion of the *go - went* type is listed (see e.g. Lieber 1981). This, however, does not capture the fact that a number of the vowel patterns are recurrent and seem to form family resemblance clusters. Bybee and Newman (1994) have therefore argued that stem changes (as well as inflectional affixes) are represented in the lexicon via associative pattern networks or schemas; see Köpcke (1998) for an attempt to classify some of the strong verbs of German in terms of schemas. A third variant of a lexicalist analysis of stem alternants comes from Wunderlich (1996). In this account, the morphological relationships between the various stem variants of strong verbs are represented in terms of non-monotonic default inheritance hierarchies in which stem variants constitute subnodes of hierarchically structured lexical entries. Each subnode is defined in terms of a phonological string and a morphological feature. Consider, for illustration, the lexical entry for the German verb *werfen* ‘to throw’ from Wunderlich (1996: 96):

(3)



Each node in a structured lexical entry represents a pair (<phonological string, morphological feature value>), e.g. [.i.i.] vs. [..a..] and [-1] vs. [+PRET] for *wirf-* vs. *warf-*, and each subnode inherits all information from its mother, except for the features it replaces or adds; for example, the subnode [..a..]_{+PRET} inherits the onset *w-*, the coda *-rf*, and the categorial feature [+V] from the higher node. The base or default stem form is *werf-*, the form that occurs in most present tense forms and the infinitive. The other stem variants occur under specific circumstances, e.g. *wirf-* for 2nd and 3rd sg. present tense forms and in imperatives, *warf-* in preterite forms, *(ge)worfen* in participles and *würf-* in subjunctives. Note that most stem variants have impoverished entries (to avoid redundancies) and that the various stem forms are hierarchically structured. Hence, structured lexical entries may form lexical templates, based on shared subnodes and shared structure. The subnodes of the entry in (3), for example, are shared by several other strong verbs (*sterben* ‘to die’, *verderben* ‘to spoil’, *helfen* ‘to help’). In this way, structured lexical entries are maximally underspecified, and, at

the same time, account for the overall similarity of stem alternants to their base forms and for the family resemblance structure of strong verbs.

In the following two sections, we will determine the generalization properties of different stem forms by examining the kinds of stem overgeneralization errors produced by children learning German as their native language and by testing experimentally which stem forms adult native speakers apply to nonce words. Both studies provide support for a lexicalist treatment of strong verb stems along the lines of Wunderlich (1996).

3. Stem formation errors in German child language

One important psycholinguistic source for determining the generalization properties of a linguistic pattern or rule comes from overgeneralization errors produced by children. Consider, for example, past tense formation in English child language where only one type of inflectional error was found: overapplications of the regular past tense affix *-ed* to irregular stems (**go-ed*, etc.). Overapplications of irregular patterns (**brang*, **wope*) are practically non-existent (see e.g. Marcus et al. 1992, Xu & Pinker 1995). It was also found that irregular verbs are sensitive to frequency and similarity: children make overgeneralization errors more often with low-frequency irregular verbs, and they produce fewer of them with irregular verbs that fall into families with more numerous and higher-frequency members. The same pattern of errors was found in two inflectional systems of German, past participles and noun plurals (Clahsen et al. 1992, Clahsen & Rothweiler 1993, Weyerts & Clahsen 1994, Bartke et al. 1996).

These findings have been taken to indicate that regular and irregular inflectional processes are dissociated in children's grammars in basically the same way as in the adult grammar. From a dual-mechanism perspective, the observed differences in the generalization properties of regular and irregular inflection are claimed to follow from different representations, rules (or equivalent operations) with unrestricted productivity for regular inflection, and memory-based representations that only allow for restricted (similarity-based) generalizations for irregulars. In this account, past-tense overregularization errors such as **bring-ed* are due to the child's applying a regular (*-ed*) rule in cases in which the lexical entry for the irregular word form *brought* is not available, and they disappear once the child can reliably retrieve the correct irregular word form.

For the present paper, we investigated stem forms of strong verbs in longitudinal data from 7 children and in an additional elicited production experiment with 26 children. We examined two families of strong verbs which both require stem changes in (particular kinds of) present tense forms, namely verbs of the *geben* type ('to give'), which have an *-i-* stem, and verbs of the *schlafen* type ('to sleep'), which have an unlauded stem, both occurring in 2nd / 3rd sg. present tense forms (e.g. *gib-st* 'give-2nd sg. pres.', *schlief-st* 'sleep-2nd sg. pres.');

occurs in imperatives. Other families of strong verbs require stem changes in the past participle, the preterite and/or the subjunctive. Participle formation in German children was examined in two earlier studies (Clahsen & Rothweiler 1993, Weyerts & Clahsen 1994), and the relevant results on stem formation will be summarized below. Simple preterite and subjunctive forms of main lexical verbs, on the other hand, are hardly ever used in spoken German, which is why instances of such verb forms are basically absent, or at least extremely rare, in the speech of young children.

Method

The longitudinal data we have investigated consist of 73 samples of spontaneous speech from 7 children covering the age period of 1;11 to 3;8 (see Table 1). The corpora were collected in the LEXLERN project at the University of Düsseldorf (see Clahsen et al. 1993, 1994, 1996). In these data, we extracted by hand all correct and incorrect stem forms of strong verbs of the *geben* type produced by the children in contexts that require an *-i-* stem in the adult language and of the *schlafen* type in contexts that require an unlauded stem in the adult language.

Table 1: Data

Child	Age	Number of recordings
Annelie	2;4 – 2;9	6
Hannah	2;0 – 2;7	8
Katrin	2;1 – 2;6	15
Leonie	1;11 – 2;11	15
Philipp	3;1 – 3;8	8
Sabrina	1;11 – 2;2	6
Svenja	2;9 – 3;3	15
Totals	1;11 – 3;8	73

To investigate stem formation in older children, we performed an auditory elicited production task with 26 children covering the age range of 6;2 to 10;5. The materials were recorded as wav-files and presented to the children from a laptop computer; their responses were recorded on a digital tape. The children first listened to a sentence containing an infinitive form of a verb as the final word, e.g. *sehen* in (4a). They then heard a second sentence in which the appropriate finite verb form of the previously presented verb was replaced by a beep, and they were asked to produce the correct verb form, e.g. *sieht* in (4b). In all cases, the finite verb to be filled in was the 3rd sg. present tense.

- (4) a. Martin will unbedingt den neuen Pokemon-Film *sehen*.
 ‘M. definitely wants to see the new pokemon movie’.
- b. Also gibt ihm seine Mutter Geld und Martin ^{beep} den Film. → *sieht*
 ‘Hence his mother gives him some money, and Martin ___ the movie’

There were 49 pairs of sentences such as those in (4). Five pairs preceded the actual experiment to familiarize children with the task. Of the remaining 44 sentence pairs, 22 contained the experimental items, and 22 were used as fillers. Sentences containing experimental items were pseudo-randomized with sentences containing fillers. The experimental items were 12 strong verbs of the *geben* type, and 10 strong verbs of the *schlafen* type. To examine effects of frequency, each of these two types of strong verbs was further divided into two conditions, a high-frequency and a low-frequency one, each according to the frequencies of the marked stems (i.e. the *-i-* or the umlauted stem) in the CELEX database of spoken German (Baayen et al. 1993). For verbs with *-i-* stems, the high-frequency condition consisted of 6 verbs with a mean (*-i-*) stem frequency of 37 per million, while in the low-frequency condition there were 6 verbs with a mean (*-i-*) stem frequency of 0.1 per million; for verbs that require umlauted stems, the high-frequency condition consisted of 5 verbs with a mean (umlaut) stem frequency of 6.1 per million, while in the low-frequency condition there were 5 verbs with a mean (umlaut) stem frequency of 0.1 per million.

Results

Consider first the results from the longitudinal data of the younger children. Here we found two types of error. The first one is overapplication of the unmarked (non-alternated) stem in cases in which strong stem forms are required in the adult language. Some examples of such stem overregularizations are given in (5).

- (5) a. **lauft* (= *läuft*) ‘run’ Katrin (2;4)
 b. **lest* (= *liest*) ‘read’ Svenja (3;2)

There are 84 stem errors of this kind in the longitudinal data against 122 strong stem forms with correct *-i-* stems for verbs of the *geben* type or correct umlauted stems for verbs of the *schlafen* type. Note that (as illustrated by the examples in (5)) most stem errors co-occur with the correct person and number suffix; there are only 11 cases that might be analysed as bare uninflected forms or infinitives, e.g. *der fang* ‘he catch’ (Svenja, 3;2) or *du nehmen* ‘you take’ (Sabrina, 2;0). Not counting these 11 cases, we are left with 73 stem overregularizations for verb forms that are clearly finite (in terms of their suffixes) and in which a strong stem form is required in the adult language.

The second type of error is incorrect occurrences of marked *-i-* or umlauted forms such as those illustrated in (6):

- (6) a. *alle fällt da runter* (correct: *fall-en*) Katrin (2;5)
 ‘everybody fall down there’
 b. *ich gib dir das* (correct: *geb(e)*) Philipp (3;4)
 ‘I give you that’
 c. *ich sieh* (correct: *seh(e)*) Svenja (3;1)
 ‘I see’

This error type is much less common than overapplications of the unmarked stems. There were 10 cases such as those in (6) in the entire longitudinal data set.

Interestingly, even though these 10 forms are incorrect in the syntactic contexts in which the children use them, the stem forms produced by the children are not incorrect. For example, the forms *fäll-*, *gib-* and *sieh-* are the correct adult stem forms for 2nd and 3rd sg. present tense indicative forms of these verbs, i.e. *fallen*, *geben* and *sehen*. The errors in the children's utterances arise because the children apply a marked stem form to a 1st sg. form and (in one case) to a 3rd pl. form. That is, a particular morpho-syntactic constraint on the use of marked stem forms, namely that *-i-* stems and unlauded stems are restricted to 2nd/3rd sg. (and to imperatives) is violated. Note, however, that there was no single instance of an irregularization error, i.e. a case in which a marked (alternated) stem would be overapplied to a weak verb. Thus, the distribution of stem errors in the longitudinal data shows a clear contrast between unmarked and strong stems; the former are frequently overregularized, in 39.7% (i.e. 73 out of 184 cases), whereas stem irregularizations (i.e. strong stems overapplied to weak verbs) are non-existent.

Another finding concerns the *development* of the stem overregularization errors over time. To investigate developmental changes, the various recordings were assigned to stages of development defined in terms of the mean length of utterance (MLU); see Brown (1973). Table 2 shows a breakdown of the stem overregularization errors against the number of correct marked stem forms across these stages.

Table 2: Development of stem overregularization errors

Stages	Number of stem errors	Number of correct stems	Error percentages
I (MLU 1.75)	1	11	8.3%
II (1.75 < MLU 2.75)	30	46	39.4%
III (2.75 < MLU 3.5)	8	17	32.0%
IV (MLU > 3.5)	34	48	41.4%
Totals	73	122	37.4%

Table 2 shows that the period from stage II onwards in which stem overregularization errors frequently occur is preceded by a period (= Stage I) in which almost all of the required marked stem forms are correctly produced by the children. Moreover we can see from Table 2 that in the age period represented in the longitudinal data, there are no signs of stem overregularization errors disappearing from the speech of the children or decreasing over time.

Consider now the experimental results from the older children. We found that the children did not show any difficulty performing the task: 97% of their responses were appropriate in that they produced a finite form of the prompted lexeme, e.g. a finite form of *sehen* in (4). There are, however, many stem errors, even in these older children, all of which were overregularizations of the unmarked (non-alternated) stem in cases in which strong stem forms are required in the adult language, i.e. errors such as those in (5) above. Table 3 presents a breakdown for the experimental conditions.

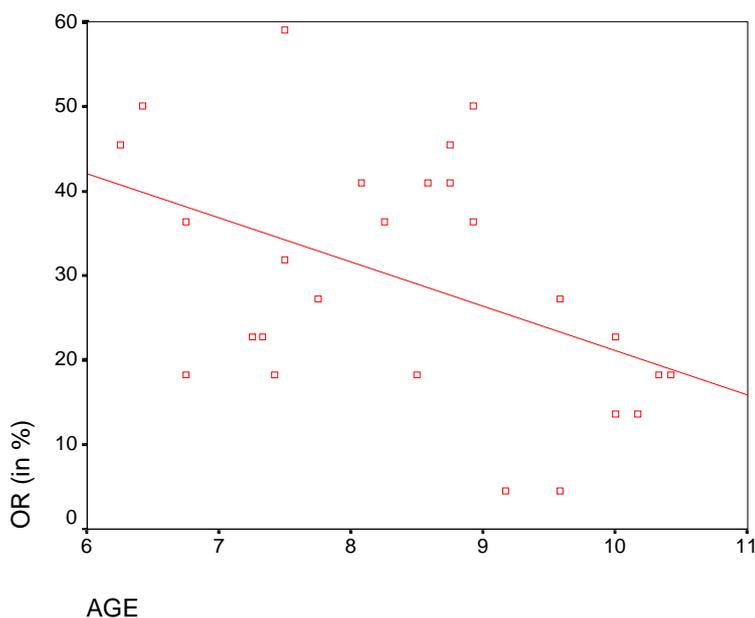
Table 3: Elicited stem forms

Required stem form and stem frequency	Number of stem errors	Number of correct stems	Error percentages
-i- / high frequency	19	136	12.2%
-i- / low frequency	89	59	37.3%
-ä- / high frequency	9	118	7.0%
-ä- / low frequency	51	74	40.8%
Totals	168	387	30.2%

Table 3 shows a clear frequency effect in the stem overregularizations for both classes of strong verbs: low-frequency stems elicit significantly more stem errors than high-frequency ones ($t(25) = 10.399, p < .001$).

Finally, in order to examine developmental changes, we plotted in Figure 1 the percentages of overregularizations against age. A regression analysis revealed that stem overregularizations are significantly linked to age and that the overregularization errors gradually decrease with age ($R^2 = .174, f(x) = 73.51 - 5.243x, F(1,24) = 6.252, p < .05$).

Figure 1: Stem overregularizations in relation to age



Discussion

The results on children's stem formations obtained in the present study are similar to those of earlier studies on participle forms of strong verbs (Clahsen & Rothweiler 1993, Weyerts & Clahsen 1994). In these studies, we examined longitudinal data from 9 children between the ages of 1;4 and 3;9 and elicited production data from 70

children between ages 3;6 and 8;10; there were 222 stem errors in participle forms, more than 90% of which were cases of unmarked (non-alternated) stems in cases in which strong stem forms are required in the adult language, e.g. **gehelft* instead of the correct form *geholfen* ‘helped’. We also found that the period in which children produce stem overregularizations is preceded by a stage without stem errors in their participle forms, and that stem overregularizations do not disappear during the age period under study; in the elicited data, even the 7- to 8-year olds produced such errors. These results are parallel to what we found for the two families of strong verbs examined in the present study.

The results obtained on children’s stem formation errors provide support for the view that stem alternants are stored in the permanent lexicon, rather than being derived by rule. Consider, for example, the frequency effect found for stem overregularizations; low-frequency forms yielded significantly more stem errors than high-frequency ones. Memory storage and retrieval is dependent on frequency of exposure, and the frequency effect obtained for strong stems suggests that they are stored in lexical memory. Additional evidence for this comes from the fact that the number of errors gradually decreases with age; this is because when children get older it is likely that memory traces for the correct irregulars are becoming stronger and the children’s ability to retrieve them is becoming more reliable.

One way of accounting for the kinds of stem error produced by children is in terms of structured lexical entries such as those proposed by Wunderlich (1996) for strong verbs of German (see (3) above). Recall that in such representations, the various stem alternants are not listed separately and completely, but that strong stems may have impoverished (i.e. underspecified) entries and that when these items are used, their base entry is filled in to obtain the full interpretation and form. Given this format, stem errors arise when subnode information containing the correct pairing of phonological strings and morphological feature values for a strong stem form is not available or not accessible to children. In such cases, they fall back on the base entry, i.e. the highest form of a structured lexical entry, producing errors such as **helf-st* instead of the correct *hilf-st*. In this way, the unmarked base stem serves as a default form in circumstances in which the required specific forms are not retrieved.

4. Generalizing stem forms in adult native speakers

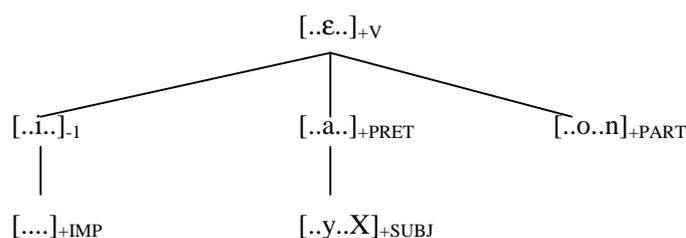
Another method for determining the generalization properties of inflectional patterns or rules are elicited production experiments with nonce words. In such experiments, participants are presented with one or two forms of a nonce word, they are then asked to repeat the nonce form, and finally they have to produce a different inflectional form of that nonce word which they have not seen before. Such experiments have been used

to study the generalizability of different inflectional phenomena (see e.g. Bybee & Pardo 1981, Clahsen 1997).

For the present study, we examined strong verbs such as *werfen* ‘to throw’ which have an *-e-* stem as the unmarked base and an *-i-* stem for imperatives and 2nd/3rd sg. present tense forms (see (3) above). Two observations are relevant for the distribution of these stem forms. The first one is that strong verbs with an *-e-* stem in the infinitive and an *-i-* stem in the imperative always exhibit the *-i-* stem in 2nd/3rd sg. present tense forms; compare *geben - gib - gibst* (*gebst), ‘to give - give-imp. - give-2nd sg.’. The second observation is that most verbs with *-e-* stems in the infinitive and an *-i-* stem in 2nd/3rd sg. present tense forms also have an *-i-* stem in imperative forms; there are, however, some exceptions, e.g. *werden - werd(e)* (*wird) - *wirst* ‘to become - become-imp. - become-2nd sg.’.

These two generalizations can be captured in different ways. One possibility would be to postulate a rule that derives *-i-* forms from *-e-* stems, for both imperative and 2nd/3rd sg. present tense forms of strong verbs. A second possibility would be to assume that both *-e-* and *-i-* stems are stored separately and that the patterns mentioned above follow from associations between them. A third possibility would be to account for the distribution of these stem forms in terms of lexical templates (Wunderlich 1996). Consider, for example, the lexical template in (7) which is derived from the structured lexical entries of strong verbs such as *werfen* ‘to throw’, *sterben* ‘to die’, *verderben* ‘to spoil’. Here, the subnodes for imperatives and 2nd/3rd sg. present tense forms are asymmetrically connected to each other, such that the imperative node fully inherits its stem form from the 2nd/3rd sg. present tense forms, whereas the onsets and the codas of the stems of 2nd/3rd sg. present tense forms are inherited from the unmarked base form and the *-i-* stem vowel change is introduced at this level. This is meant to capture the difference mentioned above in the distribution of *-i-* stems between imperatives and 2nd/3rd sg. present tense forms.

(7)



From this account, we would expect to find differences in generalizability between *-e-* and *-i-* stems. The former represents the unmarked base stems and should therefore be used under default circumstances, i.e. when it is not blocked by a more specific entry; *-i-* stems should be generalized more restrictively, i.e. only to strong verbs which have subnodes with marked stem forms. Moreover, in (7) the imperative stem is fully inherited from the stem of a 2nd/3rd sg. present tense form, while the latter is inherited from the unmarked base and involves a stem vowel change. We would expect to find

corresponding experimental differences between imperatives and 2nd/3rd sg. present tense forms in the way subjects are willing to generalize *-i-* stems.

Method

36 adult native speakers of German (mean age: 33) participated in a paper and pencil task in which they were asked to complete sentences in a booklet by filling in blanks. The booklet consisted of 60 different short stories containing phonotactically well-formed nonce verbs and nouns. The crucial experimental items were the nonce verbs; nonce nouns were added to reduce uncontrolled semantic associations. The experiment consisted of three steps. Subjects were first given each nonce verb in the infinitive and the imperative or 2nd/3rd sg. present tense form ('step 1'). They were then asked to use the previously presented forms to fill two blanks, one for the infinitive and one for imperative or 2nd/3rd sg. present tense form ('step 2'). Finally, they had to fill in a third blank, this time by providing a form of the nonce verb that was not presented to them before ('step 3'); see (8) for illustration.

There were three sets of short stories. In one set of 20 stories, the nonce verbs were introduced in their infinitive and a 2nd/3rd sg. present tense form. In step 2, participants had to supply both of these verb forms, and in the third (crucial) step, they had to fill in the imperative form of the nonce verb. In 10 of these stories, the 2nd/3rd sg. present tense form had the unmarked base stem (*-e-*); in the other 10 stories, the 2nd/3rd sg. form was presented with the strong *-i-* stem. In the second set of 20 stories, participants were given the infinitive and an imperative form of a nonce verb, and they had to fill in the 2nd/3rd sg. present tense form in step 3. In 10 of these stories, the imperative had the unmarked *-e-* stem, in the remaining 10 stories the imperative was presented with the strong *-i-* stem.

(8) *Step 1: presentation of nonce verb in infinitive and imperative or 2nd/3rd sg. present tense*

- a. Peter hat meistens keine Lust, den Prum zu *quelmen*.
'Most of the time, Peter is not keen to *quelm* the prum.'
- b. Wenn er ihn doch einmal *quelmt/quilmt*, hat er noch drei Tage später schlechte Laune.
'If he *quelms/quilms* it anyway, he is still in a bad mood even three days later.'

Step 2: repetition of verb forms presented

- c. Wenn es wieder an der Zeit ist, den Prum zu → *quelmen*,
'If it is time to *quelm* the prum again'
- d. muß Susi dafür sorgen, daß Peter ihn → *quelmt/quilmt*.
'Susi must take care that Peter *quelms/quilms* it.'

Step 3: production of novel verb form

- e. Meistens muß sie Peter anschreien: Verdammt noch mal → *quelm/quilm* endlich mal den Prum!

‘Most of the time, she has to yell at Peter: “Damn it, *quelm/quilm* the prum!”’

In addition to the 40 experimental items mentioned above, participants were given 20 other short stories in which they were presented with sentences containing infinitives and regular or strong preterite forms of nonce verbs; they were asked to repeat these forms and to produce the corresponding past subjunctive form. Filler items were added to obscure similarities between experimental items and to prevent subjects from developing any particular task-specific response strategies. All experimental and the filler items were pseudo-randomized.

Results

For the following analysis, we only included the participants’ responses of a given verb in step 3 if they had correctly reproduced the infinitive and the imperative or 2nd/3rd sg. forms of that verb in step 2; this was the case in 94% of the participants’ responses, i.e. in 1352 out of 1440 cases. Table 4 presents a breakdown of the participants’ responses in step 3 (shown in percentages), separately for verbs that were presented with an *-e-* stem in their imperative or 2nd/3rd sg. forms in step 1, and for verbs that were presented with an *-i-* stem in their imperative or 2nd/3rd sg. forms in step 1; absolute numbers of responses for each category are shown in brackets.

Table 4: Stem forms of imperative and 2nd/3rd sg. present tense forms of nonce verbs

	<i>Participants’ responses in step 3</i>							
	imperative				2 nd /3 rd sg. present tense			
<i>Imp. or 2nd/3rd sg. presented in step 1</i>	<i>-e-</i>	<i>-i-</i>	other vowel	other form	<i>-e-</i>	<i>-i-</i>	other vowel	other form
with <i>-e-</i> stem	89% (303)	9% (32)	0% (0)	1% (4)	88% (300)	6% (21)	0% (0)	6% (21)
with <i>-i-</i> stem	15% (51)	83% (280)	0.3% (1)	2% (6)	22% (73)	71% (235)	1% (2)	7% (23)

Table 4 shows that in more than 90% of the newly created verb forms subjects made use of *-e-* or *-i-* stems. There were only three instances with correct inflectional endings, but with the stem vowel *-a-*. There was also a small number of cases in which ‘other forms’, e.g. past tense forms were produced. Table 4 also shows that participants preferred the vowel with which the novel verb was introduced to them in step 1. However, this preference was stronger for *-e-* than for *-i-* stems. In more than 88% of all stories, the unmarked *-e-* stem presented in step 1 was also employed by the participants in their own word formations in step 3. In contrast, the strong *-i-* stems were used in 83% of the imperative contexts and 71% of the 2nd/3rd sg. present tense contexts. Instead of maintaining the *-i-* stem from step 1, participants relied on the *-e-* stem. Finally, we note that the *-i-* stem is more often used for imperatives (83%, 9%) than for 2nd/3rd sg. present tense forms (71%, 6%).

The observed differences are also confirmed statistically. We determined the number of cases in which in step 3 participants made use of the stem form that was presented to them in step 1. For these responses, we carried out an Analysis of Variance with two factors, *Stem Type* (-e- vs. -i-) and *Verb Form* (imperative vs. 2nd/3rd sg.). This analysis revealed the following effects. First, a significant main effect of *Stem Type* which is due to the overall preference for -e- stems ($F_1(1,35)=12.048$, $p<.002$; $F_2(1,59)=39.064$, $p<.001$). Second, a significant main effect of *Verb Form* ($F_1(1,35)=9.492$, $p<.005$; $F_2(1,59)=8.281$, $p<.007$), due to the fact that participants were more likely to maintain a stem form from step 1 when they had to form an imperative in step 3 than a 2nd/3rd sg. form. Finally, there was a significant *Stem Type X Verb Form* interaction ($F_1(1,35)=4.777$, $p<.005$; $F_2(1,59)=6.138$, $p<.02$). A subsequent pairwise comparison revealed that this is due to participants producing -i- stems more often for imperatives than for 2nd/3rd sg. forms (83% vs. 71%, $t(35)=3.5$, $p<.002$ (subject analysis); $t(59)=3.7$, $p<.002$ (item analysis)).

Discussion

The first finding from the present experiment is an asymmetry between the generalization properties of -e- and -i- stems in the subjects' responses: -e- was frequently used to form imperative and 2nd/3rd sg. present tense forms, even when the verb had been introduced as a strong verb (i.e. with an -i- stem). Yet, -i- stems were hardly ever generalized to nonce verbs that were introduced with -e- stems. This finding confirms the results from the acquisition study, in which children were shown to overregularize the unmarked base stem (-e-) to contexts in which a strong -i- stem would have been required, but not vice versa. These results show that the unmarked base stem serves as a default form and generalizes widely, while -i- stems are used more restrictively, both by children and adults. On the other hand, it was not the case that -i- stems were absent from the participants' formations of imperative and 2nd/3rd sg. forms in the present experiment. Indeed, when a nonce verb with an -i- stem was presented to them, participants were more likely to employ this stem in their imperative or 2nd/3rd sg. formations than to rely on the unmarked base stem.

As mentioned above, there are three possibilities of how one may attempt to explain these findings, in terms of (a) rules for marked stem forms, (b) associative connections between different forms, and in terms of (c) the hierarchically structured lexical template in (7). It is not clear how accounts (a) and (b) would explain the finding that participants produced a given -i- stem significantly more often when they were asked to form an imperative than when they had to produce a present tense form. Questions also remain as to how the rule-based and the associative models could explain the observed asymmetry between -e- and -i- stems in their generalization properties.

The third view, on the other hand, according to which the different stem forms of strong verbs are represented in lexical templates such as (7) does seem to provide an account for our experimental findings. The unmarked base stem is represented as the highest node in (7); it is employed by default in cases in which it is not blocked by a

more specific stem form further down on the inheritance hierarchy, and hence the wide generalizability of *-e-* stems in our child and adult experiments. Another relevant property of the lexical template in (7) is that the two subnodes for *-i-* stems, the one for imperatives and the one for 2nd/3rd sg. present tense forms, are asymmetrically represented on different levels of the inheritance hierarchy. While the imperative stem is fully inherited from the 2nd/3rd sg. present tense stem, the latter is directly connected to the unmarked base stem and inherits all its features except the *-i-* vowel change and the morphological feature [-1] specifying 2nd/3rd sg. This is compatible with the observed asymmetry in the adult experiment. Recall that participants were significantly more likely to supply an *-i-* stem when their task was to form an imperative after having been presented with an *-i-* stem of a 2nd/3rd sg. present tense form (e.g. *quelmen - (er) quilmt* → *quilm!*, see (8) above) than vice versa. Given the structure of the lexical template, the imperative stem is directly inherited from the 2nd/3rd sg. form presented to the participants. In contrast, when their task is to produce a 2nd/3rd sg. present tense form, inheritance is from the unmarked base stem, and instead of full inheritance of the stem form only the onset and the coda of the base stem are inherited to the stem of the 2nd/3rd sg.; the *-i-* stem vowel change is newly introduced at the subnode. Hence, the observed differences between the use of *-i-* stems on imperatives and on 2nd/3rd sg. as well as the productivity of *-e-* stems provide support for the default inheritance structure of lexical templates such as (7).

5. Stem forms in word recognition

While the experiments presented in the previous sections examined the *production* of strong stem forms, we investigated in a separate study (Clahsen et al. 2001) their behaviour in on-line word *recognition* experiments. Here, we present a brief summary of these experimental results.

Two experiments from Clahsen et al. (2001) on strong stems employed a visual word/non-word discrimination task ('lexical decision') with reaction time (RT) as the dependent variable. Lexical decision times on non-inflected simplex words have consistently been shown to be affected by word frequency: subjects take less time to decide that high-frequency items are existing words than they do for low-frequency items (see Balota 1994 for review). This is conceived of as a memory effect: as memory traces get stronger with additional exposure, high-frequency entries can be more readily accessed than low-frequency ones. We used this task to examine whether strong stem forms exhibit the same kinds of memory-based frequency effect as uninflected lexical items. If strong stems are indeed lexically represented (rather than rule-based), there should be such effects, and high-frequency stems should produce shorter RTs in lexical decision than low-frequency ones. We also examined whether strong stems are separately represented from the inflectional suffixes with which they may co-occur. The results were as follows.

In the first lexical decision experiment, we obtained a stem frequency effect for strong preterite stem forms that were suffixed with regular person and number affixes, such as *sangen* ‘sang-1st/3rd pl.’. Response times for verb forms with high preterite stem frequencies (mean: 22.5 per million) were significantly shorter (mean difference: 49 msec.) than lexical decision times for verbs with relatively low preterite stem frequencies (mean: 13.1 per million). This difference was found despite the fact that the mean verb and word form frequencies and the form of the inflectional ending were held constant in the two experimental conditions. The experimental effect indicates that strong stems are lexically represented (rather than rule-based) and that stems are represented separately from the inflectional affixes with which they may occur (rather than being stored as wholes).

In the second lexical decision experiment, we found a parallel preterite stem frequency effect for A-B-B verbs such as *lügen* ‘to lie’, which exhibit the same stem change in preterite and participle forms (*log-* / *ge-log-en*). Reaction times for A-B-B verbs with high preterite stem frequencies (mean: 27.5 per million) were significantly shorter (mean difference: 66 msec.) than for those with low preterite stem frequencies (mean: 16.8 per million). These RT differences cannot be attributed to the *phonological* form of the stem forms, as the B-stem frequencies (e.g. *-log-*) were held constant in both conditions. Neither can they be explained in terms of verb or word form frequencies, as these were held constant as well. Instead, the observed response time difference appears to be linked to the different preterite stem frequencies thus suggesting that preterite stems are stored separately from participle forms despite their phonological overlap. This finding indicates that *morpho-syntactic* features form part of the lexical representations of marked stem forms.

While response times in lexical decision tasks are sensitive to the frequency of lexical entries, results from priming tasks provide more direct psycholinguistic measures for examining *relationships* between lexical entries or subentries. In priming tasks, stimuli are presented to subjects in a way that allows the researcher to control for semantic, phonological and/or morphological relations between pairs of items. Priming effects were found in a number of studies on morphological processing (see Sonnenstuhl et al. 1999 for review). In Clahsen et al. (2001), we made use of the priming technique to examine relationships between the different stem forms of strong verbs in German. If strong verbs are represented in hierarchically structured lexical entries, one would expect to find corresponding priming asymmetries determined by the stems’ feature specifications, for example, preterite stems (*warf-*) should be less effectively primable by present tense forms (*werf-*) than vice versa.

We used the cross-modal immediate repetition priming technique in which subjects hear a *spoken* prime immediately followed by a *visually* presented target form to which they make a word/non-word (lexical) decision. There were four conditions, two experimental conditions with prime-target pairs such as *warf*→*werft* ‘threw-2nd pl.→throw-2nd pl.’ vs. *werft*→*warf*, and two corresponding control conditions in which prime and target were identical, e.g. *werft*→*werft* and *warf*→*warf*. We found

that the two control conditions (in which primes and targets were identical) produced significantly shorter response times than the two experimental priming conditions (in which primes and targets were different). This effect was expected and is due to the immediate repetition of the same word form. More interesting was another finding, namely that in prime-target pairs such as *werft*→*warft* (with verb forms containing the unmarked stem as a prime and the strong stem as a target) the participants' response times to the visual targets were 77 msec. longer than in the corresponding control condition, whereas in prime-target pairs such as *warft*→*werft*, they were only 25 msec. longer; statistically, this difference turned out to be significant ($p < 0.02$). This confirms the predicted asymmetric priming pattern: preterite stems prime the unmarked base better than vice versa. The results are explainable in terms of the morpho-syntactic feature specifications of the stem forms involved. When a marked stem such as *warf-* is presented as the visual target preceded by an instance of the unmarked stem (*werf-*), the target contains a feature (= [+PRET]) that is unavailable from the prime, and this unprimed feature produced the increased target response times in the experiment. On the other hand, when an unmarked stem (*werf-*) is the target preceded by a marked stem (*warf-*), the target does not contain any unprimed features, and hence there were significantly shorter target response times than for preterite targets.

In sum, the three word-recognition experiments reported above provide further support for the view that the different stem forms of strong verbs are represented in hierarchically structured lexical entries such as (3). The stem frequency effects obtained in the two lexical decision experiments indicate that marked stem forms are accessed separately from the inflectional affixes with which they may occur and that marked stem forms are stored in the permanent lexicon. In addition, the frequency effect for preterite stems of A-B-B verbs indicates that these stems are represented separately from participle forms despite their phonological overlap. The observed priming patterns can also be explained in terms of structured lexical entries. In these entries, preterite stems contain a positively specified tense feature ([+PRET]) and are represented further down on the inheritance hierarchy than unmarked base stems which do not have a tense specification, even though they are used in present tense forms. This difference reflects the common linguistic assumptions that past (not present) is the marked tense feature in German and that unmarked feature values are lexically unspecified. The priming results correspond to these representational differences in that (tense)-marked stems turned out to be more difficult to prime than unmarked stems. The priming experiment also supports the view that the various stem forms of a strong verb are based on underspecified entries with impoverished, minimally specified feature contents. Suppose that, contrary to lexical representations such as (3) and (7), all stem forms were fully specified, for example present tense forms as [-Pret] and preterite forms as [+Pret]. In that case, our priming results would be left unexplained, because there would be an unprimed tense feature in the targets of both prime-target pairs; the [-Pret] feature of present tense targets could not be primed by preterite forms, and the [+Pret] feature of preterite targets could not be primed by present tense forms.

Instead, the observed priming asymmetries suggest that the mental representation of lexical forms utilizes underspecified representations.

6. Conclusion

Empirical results from different sources reported in the present study (analyses of naturalistic data from child corpora, elicited production with children and adults, visual lexical decision, cross-modal priming) provide psycholinguistic evidence for two linguistically inspired hypotheses; first, that inflected word forms that consist of strong stems and regular affixes have decomposed representations, and second, that strong stems are represented as subnodes of hierarchically structured entries with underspecified grammatical feature content. At a more general level, these findings support a dual-mechanism approach to inflection according to which two mechanisms are involved in the processing and representation of inflected word forms, inflectional rules or equivalent operations (e.g. for regular suffixation), and stored entries (e.g. for irregularly inflected word forms and for strong stems).

References

- Baayen, Harald; Piepenbrock, Richard; and van Rijn, H. (1993). *The CELEX Lexical Database (CD ROM)*. Philadelphia, PE: Linguistic Data Consortium, University of Pennsylvania.
- Balota, David A. (1994). Visual word recognition: The journey from features to meaning. In *Handbook of Psycholinguistics*, Morton Ann Gernsbacher (ed.), 303-349. San Diego, CA: Academic Press.
- Barbour, J. S. (1982). Productive and non-productive morphology: The case of the German strong verbs. *Journal of Linguistics* 18, 331-354.
- Bartke, Susanne; Marcus, Gary F.; and Clahsen, Harald (1996). Acquiring German noun plurals. In *Proceedings of the 19th Annual Boston University Conference on Language Development*, Dawn M. MacLaughlin et al. (eds.), 60-69. Boston, MA: Cascadilla Press.
- Beedham, Christopher (1994). The role of consonants in marking strong verb conjugation in German and English. *Folia Linguistica* 28, 279-296.
- (1995/96). Vowel+consonant and consonant+vowel sequences in the strong verbs of German and English. *Cahiers Ferdinand de Saussure* 49, 139-163.
- Bittner, Andreas (1996). *Starke 'schwache' Verben - schwache 'starke' Verben. Struktur des deutschen Verbsystems*. Tübingen: Stauffenburg.
- Brown, Roger (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

- Bybee, Joan L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* 10, 425-455.
- ; and Newman, Jean E. (1994). Are stem changes as natural as affixes? *Linguistics* 32, 33-34.
- ; and Pardo, Elly (1981). On lexical and morphological conditioning of alternations: A nonce-probe experiment with Spanish verbs. *Linguistics* 9, 937-968.
- Chomsky, Noam (1970). Remarks on nominalization. In *Readings in English Transformational Grammar*, Roderick A. Jacobs et al. (eds.), 184-221. Waltham, MA: Ginn.
- Clahsen, Harald (1997). The representation of participles in the German mental lexicon: Evidence for the dual-mechanism model. In *Yearbook of Morphology 1996*, Geert Booij et al. (eds.), 73-96. Dordrecht: Kluwer.
- (1999). Lexical entries and rules of language: A multi-disciplinary study of German inflection. *Behavioral and Brain Sciences* 22, 991-1060.
- ; Eisenbeiss, Sonja; Hadler, Meike; and Sonnenstuhl, Ingrid (2001). The mental representation of inflected words: An experimental study of adjectives and verbs in German. To appear in: *Language* 77.
- ; Eisenbeiss, Sonja; and Penke, Martina (1996). Lexical learning in early syntactic development. In *Generative Perspectives on Language Acquisition*, Harald Clahsen (ed.), 129-160. Amsterdam: Benjamins.
- ; Eisenbeiss, Sonja; and Vainikka, Anne (1994). The seeds of structure. A syntactic analysis of the acquisition of case marking. In *Language Acquisition Studies in Generative Grammar*, Teun Hoekstra et al. (eds.), 85-118. Amsterdam: Benjamins.
- ; Penke, Martina; and Parodi, Teresa (1993). Functional categories in early child German. *Language Acquisition* 3, 395-429.
- ; and Rothweiler, Monika (1993). Inflectional rules in children's grammars: Evidence from the development of participles in German. In *Yearbook of Morphology 1992*, Geert Booij et al. (eds.), 1-34. Dordrecht: Kluwer.
- ; Rothweiler, Monika; Woest, Andreas; and Marcus, Gary F. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition* 45, 225-255.
- Corbett, Greville G.; and Fraser, Norman M. (1993). Network Morphology: A DATR account of Russian nominal inflection. *Journal of Linguistics* 29, 113-142.
- Durrell, Martin (1980). Morphophonologische und morpholexische Regelmäßigkeiten im deutschen Ablautsystem. *Jahrbuch für internationale Germanistik* 8.2 (Reihe A), 19-28.
- (2000). Strong verb ablaut in the West Germanic languages. Unpublished ms., University of Manchester.
- Elman, Jeffrey L.; Bates, Elizabeth A.; Johnson, Mark H.; Karmiloff-Smith, Annette; Parisi, Domenico; and Plunkett, Kim (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
- Jackendoff, Ray (1975). Morphological and semantic regularities in the lexicon. *Language* 51, 639-671.

- Köpcke, Klaus-Michael (1998). Prototypisch starke und schwache Verben der deutschen Gegenwartssprache. *Germanistische Linguistik* 141/142, 45-60.
- Lieber, Rochelle (1981). *On the Organization of the Lexicon*. Bloomington, IN: Indiana University Linguistics Club.
- Marcus, Gary F.; Pinker, Steven; Ullman, Michael; Hollander, Michelle; Rosen, T. John; and Xu, Fei (1992). Overregularization in Language Acquisition. *Monographs of the Society for Research in Child Development* 57 (Serial No. 228), 1-165.
- Marslen-Wilson, William; and Tyler, Lorraine K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences* 2, 428-435.
- Pinker, Steven (1999). *Words and Rules: The Ingredients of Language*. New York, NY: Basic Books.
- Rumelhart, David E.; and McClelland, James L. (1986). On learning the past tenses of English verbs. In *Parallel Distributed Processing*, Vol. 2, James L. McClelland et al. (eds.), 216-271. Cambridge, MA: MIT Press.
- Sonnenstuhl, Ingrid; Sonja Eisenbeiss; and Harald Clahsen (1999). Morphological priming and the mental lexicon: evidence from German. *Cognition* 72, 203-236.
- Weyerts, Helga; and Clahsen, Harald (1994). Netzwerke und symbolische Regeln im Spracherwerb: Experimentelle Ergebnisse zur Entwicklung der Flexionsmorphologie. *Linguistische Berichte* 154, 430-460.
- Wiese, Richard (1996). *The Phonology of German*. Oxford: Oxford University Press.
- Wunderlich, Dieter (1996). Minimalist Morphology: The role of paradigms. In *Yearbook of Morphology 1995*, Geert Booij et al. (eds.), 93-114. Dordrecht: Kluwer.
- ; and Ray Fabri (1995). Minimalist Morphology: An approach to inflection. *Zeitschrift für Sprachwissenschaft* 14, 236-294.
- Xu, Fei; and Pinker, Steven (1995). Weird past tense forms. *Journal of Child Language* 22, 531-556.